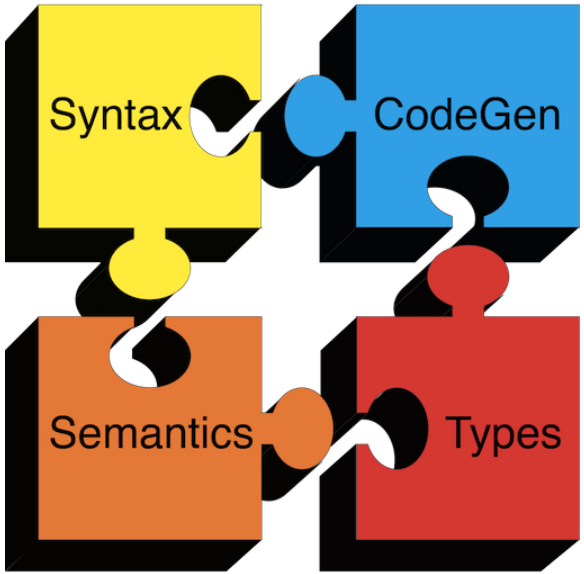

Compilers

Regular Languages

Regular Languages

- Lexical structure = token classes
- We must say what set of strings is in a token class
  - Use *regular languages*

Regular Languages

- Single character

- Epsilon

Regular Languages

- Union
- Concatenation
- Iteration

Regular Languages

- **Def.** The *regular expressions over* Σ are the smallest set of expressions including

Regular Languages

Regular Languages

Choose the regular languages that
are equivalent to the given regular language: $(0 + 1)^*1(0 + 1)^*$

- ☐ $(01 + 11)^*(0 + 1)^*$
- ☐ $(0 + 1)^*(10 + 11 + 1)(0 + 1)^*$
- ☐ $(1 + 0)^*1(1 + 0)^*$
- ☐ $(0+ 1)^*(0 + 1)(0 + 1)^*$

$\Sigma = \{ 0, 1 \}$

Regular Languages

- Regular expressions specify regular languages

- Five constructs
  - Two base cases
    - empty and 1-character strings
  - Three compound expressions
    - union, concatenation, iteration